(19) Canadian Intellectual Property Office
An Agency of Industry Canada

Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

(11) CA 2 463 166 (13) A1

(40) 24.04.2003
(43) 24.04.2003

(12)

(21) 2 463 166

(22) 11.10.2002

(51) Int. Cl.[7]: C12N 15/52, A01H 5/00, C12N 15/82

(85) 07.04.2004

(86) PCT/GB02/004642

(87) WO03/033713

(30) 0124574.5 GB 12.10.2001

(71)
BIOGEMMA UK LIMITED,
200 Cambridge Science Park
Milton Road
CB4 0GZ, CAMBRIDGESHIRE, XX (GB).

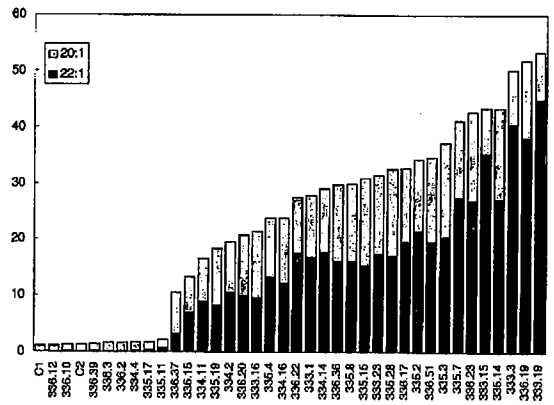
(72)
SLABAS, ANTONI RYSZARD (GB).
WALLINGTON, EMMA JANE (GB).
WILMER, JEROEN ALEXANDER (GB).

(74)
ROBIC

(54) BIOSYNTHESE D'HUILE
(54) PRODUCTION OF OIL WITH HIGHER ERUCIC ACID CONTENT

(57)
The invention provides a method of producing, in a plant, oil having an erucic acid content above 66 % erucic acid, the method comprising (i) expressing in the plant nucleic acid encoding an elongase and nucleic acid encoding an acyltransferase enzyme; and (ii) extracting oil from the plant. Corresponding plants seeds and oils are also provided.

Distribution of erucic acid levels in LEAR (mol% if fatty acids)



Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2463166 A1 2003/04/24

(21) 2 463 166

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 2002/10/11
(87) Date publication PCT/PCT Publication Date: 2003/04/24
(85) Entrée phase nationale/National Entry: 2004/04/07
(86) N° demande PCT/PCT Application No.: GB 2002/004642
(87) N° publication PCT/PCT Publication No.: 2003/033713
(30) Priorité/Priority: 2001/10/12 (0124574.5) GB

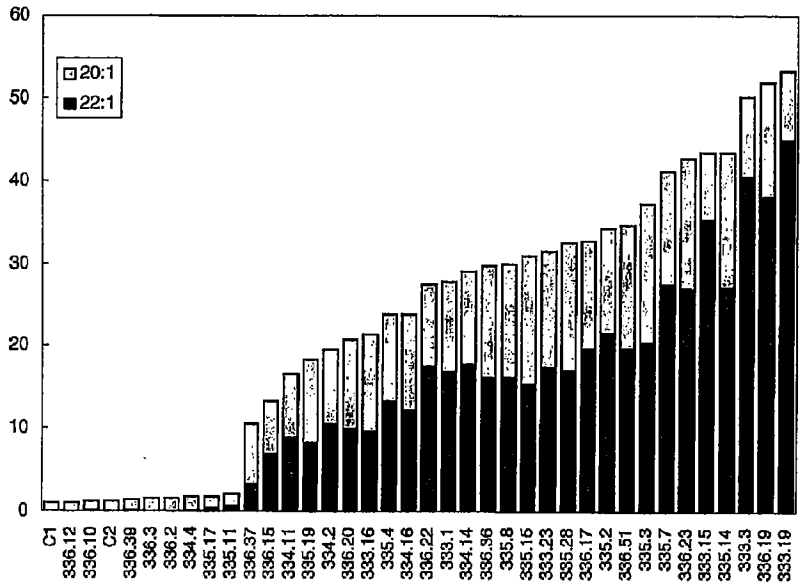
(51) Cl.Int.^{7}/Int.Cl.^{7} C12N 15/52, A01H 5/00, C12N 15/82
(71) Demandeur/Applicant:
BIOGEMMA UK LIMITED, GB

(72) Inventeurs/Inventors:
WILMER, JEROEN ALEXANDER, GB;
WALLINGTON, EMMA JANE, GB;
SLABAS, ANTONI RYSZARD, GB

(74) Agent: ROBIC

(54) Titre : BIOSYNTHESE D'HUILE
(54) Title: PRODUCTION OF OIL WITH HIGHER ERUCIC ACID CONTENT

Distribution of erucic acid levels in LEAR (mol% if fatty acids)



(57) Abrégé/Abstract:
The invention provides a method of producing, in a plant, oil having an erucic acid content above 66 % erucic acid, the method comprising (i) expressing in the plant nucleic acid encoding an elongase and nucleic acid encoding an acyltransferase enzyme; and (ii) extracting oil from the plant. Corresponding plants seeds and oils are also provided.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
24 April 2003 (24.04.2003)

PCT

(10) International Publication Number
WO 03/033713 A3

(51) International Patent Classification[7]: C12N 15/52, 15/82, A01H 5/00

(21) International Application Number: PCT/GB02/04642

(22) International Filing Date: 11 October 2002 (11.10.2002)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
0124574.5 12 October 2001 (12.10.2001) GB

(71) Applicant *(for all designated States except US)*: **BIOGEMMA UK LIMITED** [GB/GB]; 200 Cambridge Science Park, Milton Road, Cambridge CB4 0G2 (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **WILMER, Jeroen, Alexander** [DK/GB]; 200 Cambridge Science Park, Milton Road, Cambridge, Cambridgeshire CB4 0G2 (GB). **WALLINGTON, Emma, Jane** [GB/GB]; 200 Cambridge Science Park, Milton Road, Cambridge, Cambridgeshire CB4 0G2 (GB). **SLABAS, Antoni, Ryszard** [GB/GB]; Department of Biological Sciences, University of Durham, Durham DH1 3LE (GB).

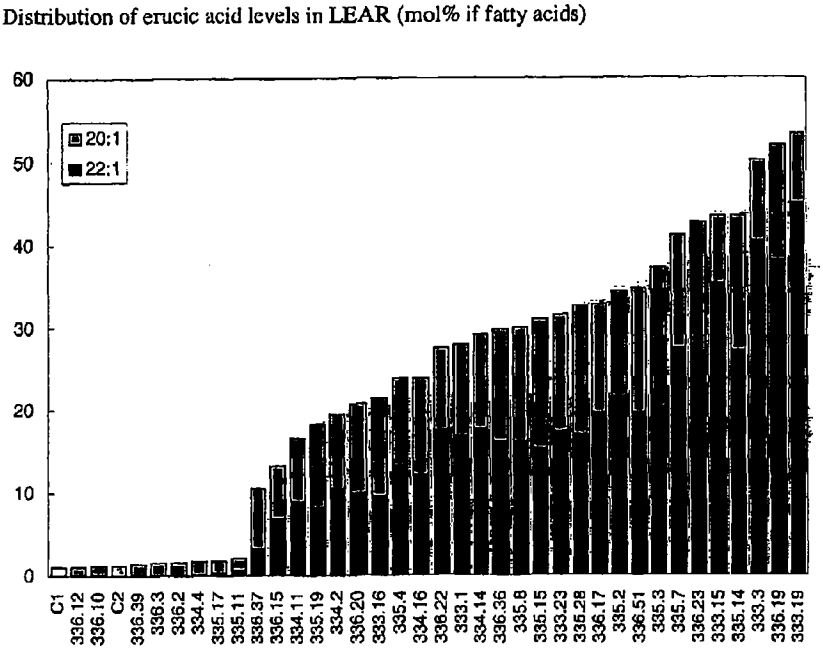
(74) Agents: **CHAPMAN, Paul, William** et al.; Kilburn & Strode, 20 Red Lion Street, London WC1R 4PJ (GB).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),

[Continued on next page]

(54) Title: PRODUCTION OF OIL WITH HIGHER ERUCIC ACID CONTENT



(57) Abstract: The invention provides a method of producing, in a plant, oil having an erucic acid content above 66 % erucic acid, the method comprising (i) expressing in the plant nucleic acid encoding an elongase and nucleic acid encoding an acyltransferase enzyme; and (ii) extracting oil from the plant. Corresponding plants seeds and oils are also provided.

WO 03/033713 A3

WO 03/033713 A3

Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report*

(88) Date of publication of the international search report:
28 August 2003

(48) Date of publication of this corrected version:
2 October 2003

(15) Information about Correction:
see PCT Gazette No. 40/2003 of 2 October 2003, Section II

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

OIL BIOSYNTHESIS

The present invention relates to the production of commercially useful oil in plants. In particular, the invention provides a method for producing such oil in plants. Also provided are the plants or parts thereof from which the oil is derived, use of the plant, and the oil itself.

Plants have long been a commercially valuable source of oil. Traditionally, plant oils were used for nutritional purposes. Recently, however, attention has focused on plant oils as sources of industrial oils, for example as replacements for, or improvements on, mineral oils. Given that oil seeds of commercially useful crops such as *Brassica napus* contain a variety of lipids (Hildish & Williams, "Chemical Composition of Natural Lipids", Chapman Hall, London, 1964), it is desirable to tailor the lipid composition to better suit our needs, for example using recombinant DNA technology (Knauf, *TIB*tech, February 1987, 40-47).

The production of commercially desirable specific oils in plants on a large scale is limited in two ways. Some plant species make oils with very high levels of essentially pure, specific fatty acids, but these species are unable to be grown in sufficient quantities and of sufficient yield to provide a commercially valuable product. Other plant species produce sufficient amounts of oil, but the oil has low levels of the specific desired fatty acids. Nevertheless, the field of oil modification in plants is wide and a number of different products have already been designed. Rape oil containing lauric acid has been marketed, and soybeans with modified levels of unsaturated fatty acids are available. In some cases the production of speciality oils seems to be straight-forward. In others, however, a number of unexpected complications have arisen which have hampered the production of plants capable of making some specific oils. For example, mutations in plant lipid synthesis genes are generally difficult to detect due to the pleiotrophic effects of mutations on plant hardiness and yield. Even if detected, proteins involved in pathways of interest have

proved difficult to isolate due to their biochemical instability. Where regulation of such proteins has been successfully altered, results generally do not coincide with expectations, presumably due to the effect of multiple converging pathways. Examples of such problems relating to the production of *Arabidopsis* producing petroselinic acid are disclosed in Ohlrogge, *13th International Symposium on Plant Lipids*, Seville, Spain: 219 & 801, (1998). Thus, there is considerable work yet to be done in achieving reliable, large-scale production of a range of commercially desirable oils.

Broadly speaking, there are two main approaches to altering the lipid content of an oil, which to date have been applied as alternatives. Firstly, plants may be modified to produce fatty acids which are foreign to the native plant. For example, rape may be modified to produce laureate which is not naturally produced by that plant. Secondly, the pattern and/or extent of incorporation of fatty acids into the glycerol backbone of a lipid may be altered.

Lipids are formed by the addition of the fatty acid moieties into the glycerol backbone by acyltransferase enzymes. There are three positions on the glycerol backbone at which fatty acids may be introduced. The acyltransferase enzymes which are specific for each position are hence referred to as 1-, 2-, and 3- acyltransferase enzymes respectively.

One of the aims of lipid engineering is to produce oils which are high in erucic (22:1) acid. Such oils are desirable for a number of reasons, in particular as replacements and/or substitutes for mineral oils, as described above. In the case of *Brassica napus* one of the most commercially important crops cultivated today, and other oil seed *Brassica* species, e.g. *Brassica juncea*, the 2-acyltransferase positively discriminates against the incorporation of erucic acid in the second position. Thus, even in those crops where erucic acid is incorporated into the first and third positions, only a

maximum of 66% of the fatty acids of the lipid can be erucic acid. These latter varieties of rape are nevertheless known as HEAR (high erucic acid rape) varieties.

It is desirable to further increase the erucic acid content of both HEAR varieties, and other useful vegetable oil crops, for example maize, sunflower, soya, mustards and linseed. Genes encoding 2-acyltransferases have been introduced into plants, in order to try to incorporate erucic acid into the second position of the glycerol backbone, with the aim of increasing the overall erucic acid content in a lipid (Brough *et al.*, *Mol Breeding* 2: 133-142 (1996)). This was successful in the re-distribution of erucic acid in the triglyceride but has not increased the overall erucic acid content of the oil. One possible reason for this is that the levels of "free" erucic acid available for incorporation into lipids in a plant are too low to support high levels of trierucin synthesis (Millar *et al.*, *The Plant J.* 12(1) 121-131 (1997)). Thus, knowledge of the factors involved in the regulation of erucic acid levels in a plant is being sought.

In this text, the terms "free" or "available" erucic acid mean erucic acid which has not been incorporated into lipid. References to the erucic acid content of oil means that which has been incorporated into lipid.

Biochemical and genetic studies have elucidated most of the pathways involved in the production of vegetable oils (Ohlrogge & Browse, Plant Cell 7: 957-970, 1995). An enzyme involved in the synthesis of fatty acids is the fatty acid elongation enzyme (FAE) complex, also referred to as an elongase. This enzyme complex is responsible for the conversion of fatty acids 18 carbons long, such as oleic (18:1) acid, to fatty acids known as very long chain fatty acids, which include erucic acid (22:1). Given its involvement in the production of erucic acid, it is apparent that the elongase plays a role in regulation of the levels of free erucic acid in a plant. Thus, it has been suggested that, in relation to *Arabidopsis*, over-expression of the FAE1 gene may assist in obtaining higher levels of free erucic acid (Millar *et al.*, *The Plant J.* **12**(1) 121-131 (1997)). Depending upon the plant species, the products of this enzyme are

C20 and C22 saturated fatty acids, utilised in wax production in leek, or C20 to C24 monounsaturated fatty acids utilised as seed storage oils in crucifers.

Over recent years, a number of ß-keto-acyl-CoA synthetase ("elongase") genes, in addition to the *Arabidopsis* FAE1 gene, have been cloned from a variety of species. Sequence data is available for elongases isolated from *Arabidopsis* (Millar & Kunst, *Plant J.* 12: 121-131, (1997)), jojoba (Lassner *et al*, *Plant Cell* 8: 281-292, (1996)), honesty (Millar & Kunst, *Plant J.* 12: 121-131, (1997)), leek (Evenson & Post-Beittenmiller, *Plant Physiol.* 109: 707-716, (1995)) and oilseed rape (Sequence: Genbank AF009563 & BNU50771). These enzymes all produce a range of other very long chain fatty acids besides erucic acid, for example fatty acid 20:1 in *Arabidopsis* and oilseed rape, fatty acids 20:0 and 22:0 in leek and fatty acid 24:1 in honesty and jojoba.

Recently, experiments have been performed on high-erucic acid rapeseed plants in which the plants were transformed with constructs encoding an acyltransferase and an elongase. However, none of the transformants were found to contain erucic acid levels greater than 60% (Han *et al.*, *Plant Mol. Bio.* 46 229-239 (2001)).

The present invention aims to overcome or ameliorate the above problems of the prior art by enabling for the first time the production of lipid with higher levels of erucic acid.

Thus, in a first aspect of the present invention, there is provided a method of producing, in a plant, oil containing more than 66% erucic acid, wherein the method comprises (i) expressing in the plant nucleic acid encoding an elongase and nucleic acid encoding an acyltransferase; and (ii) extracting oil from the plant.

The present invention is based upon the inventors' discovery that the combination of an elongase enzyme and an acyltransferase enzyme expressed in a plant can have the

surprising effect of increasing the erucic acid content of oil to above the theoretical maximum of 66%. These results mean that for the first time it has been possible to introduce erucic acid into all three positions in the glycerol backbone in at least a portion of the lipid of the oil and also cause an overall increase in the erucic acid incorporated into lipid rather than merely redistributing the erucic acid on the glycerol backbone. Thus, for the first time the invention enables the production of oil containing above 66% erucic acid in an ergonomic and agronomic fashion. The resulting oil can be used in a variety of industrial applications, such as feedstock's for surfactants, plasticisers, and surface coatings.

The nucleic acid sequences of the invention may be DNA or RNA, or any other option. The nucleic acid may be recombinant or isolated.

The elongase enzyme expressed in the plant is preferably one capable of the production of very long chain fatty acids including erucic acid, most preferably specific for the production of erucic acid. Examples of preferred elongase enzymes are *Brassica napus* FAE1-1 and FAE1-2 of Figures 1 and 2 respectively, or elongase enzymes encoded by the *B. napus* cDNA sequeneces disclosed in WO96/13582, or encoded by the *B. napus* sequences Genbank accession nos. AF009563 and BNU50771, or similar enzymes from other *Brassica* species, *Arabidopsis* (Millar & Kunst, Plant J, **12**: 121-131, (1997), jojoba (Lassner *et. al*, *Plant Cell* **8**: 281-292, (1996)), Lunaria (Millar & Kunst, *Plant J.* **12**: 121-131, (1997)) and *Nasturtium* (WO95/15387). Also included for use in the present invention are enzymes which are substantially identical in sequence to the elongases of Figure 1 and/or Figure 2, and which share the same enzyme activity as one or more of the elongases mentioned above.

The acyltransferase enzyme to be expressed in the plant may be a 1-, 2-, or 3-acyltransferase, but most preferably is a 2-acyltransferase, and is therefore capable of introducing a fatty acid into the second position of the glycerol backbone. More

preferably, the 2-acyltransferase is capable of introducing erucic acid into the glycerol backbone. Examples of suitable 2-acyltransferase enzymes include those from *Limnanthes alba* (WO95/27791), *Limnanthes douglassi* (WO96/24674 and WO96/09394). Also included are acyltransferases which are substantially identical thereto, and which share the same enzyme activity as one or more of the acyltransferases mentioned above.

In the context of the present invention the term "substantially identical" means that the sequence has at least 50% sequence identity, desirably at least 75% sequence identity and more desirably at least 90 or at least 95% sequence identity with one or more of the above mentioned sequences. In some cases the sequence identity may be 99% or above.

"% identity", as known in the art, is a measure of the relationship between two polypeptide sequences or two polynucleotide sequences, as determined by comparing their sequences. In general, the two sequences to be compared are aligned to give a maximum correlation between the sequences. The alignment of the two sequences is examined and the number of positions giving an exact amino acid or nucleotide correspondence between the two sequences determined, divided by the total length of the alignment and multiplied by 100 to give a % identity figure. This % identity figure may be determined over the whole length of the sequences to be compared, which is particularly suitable for sequences of the same or very similar length and which are highly homologous, or over shorter defined lengths, which is more suitable for sequences of unequal length or which have a lower level of homology.

Methods for comparing the identity of two or more sequences are well known in the art. Thus for instance, programs available in the Wisconsin Sequence Analysis Package, version 9.1 (Devereux J *et al*, Nucleic Acids Res. 12:387-395, 1984, available from Genetics Computer Group, Maidson, Wisconsin, USA), for example the programs BESTFIT and GAP, may be used to determine the % identity between two polynucleotides and the % identity between two polypeptide sequences.

BESTFIT uses the "local homology" algorithm of Smith and Waterman (Advances in Applied Mathematics, 2:482-489, 1981) and finds the best single region of similarity between two sequences. BESTFIT is more suited to comparing two polynucleotide or two polypeptide sequences which are dissimilar in length, the program assuming that the shorter sequence represents a portion of the longer. In comparison, GAP aligns two sequences finding a "maximum similarity" according to the algorithm of Neddleman and Wunsch (J. Mol. Biol. 48:443-354, 1970). GAP is more suited to comparing sequences which are approximately the same length and an alignment is expected over the entire length. Preferably, the parameters "Gap Weight" and "Length Weight" used in each program are 50 and 3 for polynucleotide sequences and 12 and 4 for polypeptide sequences, respectively. Preferably, % identities and similarities are determined when the two sequences being compared are optimally aligned.

Other programs for determining identity and/or similarity between sequences are also known in the art, for instance the BLAST family of programs (Altschul S.F. *et al*, J. Mol. Biol., 215:403-410, 1990, Altschul S.F. *et al*, Nucleic Acids Res., 25:289-3402, 1997, available from the National Center for Biotechnology Information (NCB), Bethesda, Maryland, USA and accessible through the home page of the NCBI at www.ncbi.nlm.nih.gov) and FASTA (Pearson W.R. and Lipman D.J., Proc. Nat. Acac. Sci., USA, 85:2444-2448, 1988, available as part of the Wisconsin Sequence Analysis Package). Preferably, the BLOSUM62 amino acid substitution matrix (Henikoff S. and Henikoff J.G., Proc. Nat. Acad. Sci., USA, 89:10915-10919, 1992) is used in polypeptide sequence comparisons including where nucleotide sequences are first translated into amino acid sequences before comparison.

Preferably, the program BESTFIT is used to determine the % identity of a query polynucleotide or a polypeptide sequence with respect to a polynucleotide or a polypeptide sequence of the present invention, the query and the reference sequence being optimally aligned and the parameters of the program set at the default value.

Preferably, the elongase and/or acyltransferase enzyme used in the present invention will not be native to the plant (i.e. the enzyme will be foreign to the plant). Thus, the nucleic acid encoding the foreign enzymes may be referred to as a transgene. The plant hosting the nucleic acid encoding the foreign enzyme(s) will therefore be transgenic. Such a plant may be readily distinguished from native plants due to the presence of foreign genetic material which would not be found in the native, non-transgenic plant, for example, vector sequences, marker genes, and multiple copies of nucleic acid encoding one or both of the above enzymes.

The transgene may encode an enzyme which is not present in the native, non-transformed plant, for example the 2-acyltransferase able to incorporate erucic acid at position 2 encoded by *lat2* is not present in native, non-transformed *Brassica napus*. In such case a transgenic plant may be distinguished from a native, non-transformed plant by an assay for the foreign enzyme activity or by the presence of the reaction product (in this example trierucin). As the *Limnanthes* gene encoding lat2 is not present in native, non-transformed *Brasica napus*, a transgenic plant may also be distinguished from a native, non-transformed plant by testing for the presence of the lat2 gene sequence e.g. by PCR or Southern blot.

Alternatively the transgene may encode an enzyme which is already found in native, non-transformed plants, for example the elongase enzymes encoded by BnFAE1-1 and BnFAE1-2 are found in non-transformed *Brassica napus* (because this is the plant from which the genes were isolated). In this case a transgenic plant may be identified by changed enzyme activity (e.g. level, timing) in comparison with a non-transformed plant. Presence of the transgene may be positively verified by PCR or hybridisation tests for the presence of nucleotide sequences which are known to be unique to the expression cassette containing the transgene.

Preferably, the nucleic acid sequences encoding the acyltransferase and/or elongase are expressed in the plant through the period of oil biosynthesis. This preferably

includes at least the period during which the fatty acids are incorporated into the glycerol backbone, i.e. triacylglycerols are produced. It may also include the period during which very long chain fatty acids are produced, the period of oil accumulation in the plant, and/or the period of any oil modification. The timing of each of these stages of oil biosynthesis may be readily determined by a person skilled in the art using time-course experiments using well established biochemical methods, such as GLC (Gas Liquid Chromatography) and HPLC ELSD (Evaporative Light Scatter Detection) analysis on triglycerides. The precise temporal expression pattern of the nucleic acid sequences may be adjusted according to their role in the oil biosynthesis pathway. For example, it may be desirable to express the elongase enzyme prior to the start of oil biosynthesis, during the period of fatty acid synthesis, in order to effectively increase the amount of "free" erucic acid prior to its incorporation into lipid. Similarly, the acyltransferase enzyme may be best expressed during the period of triacylglycerol synthesis. Depending upon the degree of overlap between the different stages of oil biosynthesis, it may be preferable for the elongase and acyltransferase enzymes to be expressed at the same time. In a most preferred embodiment, the temporal expression patterns of the foreign enzymes will mimic those of the corresponding native enzymes of the plant.

In addition to conferring temporal specificity on the foreign enzymes, it may also be desirable to express them in a spatially specific manner, for example to reduce any adverse effects in parts of the plant not involved in oil biosynthesis. In a preferred embodiment, the elongase and/or acyltransferase enzymes are expressed in the seed of the host plant.

The desired temporal and spatial specificity may be conferred by the use of appropriate regulatory sequences to drive expression of the nucleic acid sequences encoding the elongase and acyltransferase enzymes. Use of these regulatory sequences will allow a tighter control of transgene expression and improved co-ordination with oil biosynthesis in seeds. This will enable potentiation of transgene

activity whilst avoiding ectopic expression. Thus, in a preferred embodiment, the elongase and/or acyltransferase nucleic acid sequences to be expressed in the plant are under the control of one or more regulatory sequences capable of driving expression in a temporally and spatially specific manner. Depending upon the desired temporal and spatial specificity of the acyltransferase and elongase enzymes, it may be preferably to place each under the control of the same or separate regulatory sequences. Examples of suitable regulatory sequences include the oleosin promoter (Plant *et al* Plant Mol Biol 25(2) 193-205 (1994)), the 2S napin promoter (European patent No 0255278), and the FAE promoters of *Brassica napus* (Han *et al.*, Plant Mol. Biol. **46** 229-239 (2001)) and Arabidopsis (Rossak *et al.*, Plant Mol. Biol. **46** 717-725 (2001)). The regulatory sequences which drive expression of the native elongase and acyltransferase enzymes in the plant may also be used. The nucleic acid sequence to be expressed may also comprise 3' polyadenylation sequences, for example the Chalcone Synthase polyadenylation termination signal sequence.

In achieving the desired levels of erucic acid in the oil of the transgenic plant, it may be useful to overexpress the transgenes. Thus, the expression level will be higher than that of the native enzyme during its active period. This may be achieved by placing the transgenes under the control of a strong promoter, such as the napin promotes. The FAE promoters of *Brassica napus* may also be suitable for overexpression of the transgenes.

Where the elongase or acyltransferase enzymes are foreign to the plant into which they are introduced, it may be desirable to down-regulate, or disrupt the function of, the corresponding native enzymes or gene products in the plant. For this purpose, antisense sequences of the native elongase or acyltransferase genes may be used to suppress the expression of the native enzymes. The RNA transcribed from the antisense DNA will be capable of binding to, and destroying the function of, a sense version of the RNA found native in the cell, thereby disrupting its function. For example, antisense acyltransferase may be used in a plant where the native 2-

acyltransferase is not specific for very long chain fatty acids such as erucic acid. Preferably, the antisense sequences will be capable of hybridising to the native sequence under stringent conditions, as defined below.

It is not crucial for any such antisense sequence to be transcribed at the same time as the nucleic acid encoding the foreign acyltransferase and elongase. Antisense RNA will in general only bind when its sense complementary strand is present and so will only have its toxic effect at the appropriate time. Thus, any suitable plant promoter may be used, although it will preferably share the same spatial specificity as the native enzyme. Examples of suitable promoters include the napin promoter and the promoter of the lat2 gene from *limnanthes*.

In the context of the present invention "stringent conditions" are defined as those given in Plant genetic Transformation and Gene Expression: A Laboratory Manual, Ed. Draper *et al* 1988, Blackwell Scientific Publications, p252-255, modified as follows: prehybridization, hybridization and washes at 55-65°C, final washes (with 0.5x SSC, 0.1% SDS) omitted.

Alternatively, ribozyme technology may be used to disrupt the expression of the native elongase and/or acyltransferase in the plant. Ribozymes are RNA "enzymes" capable of highly specific cleavage against a given target sequence (Haseloff and Gerlarch, Nature 334 585-591 (1988)).

The nucleic acid sequences to be expressed in the plant may be introduced in the form of a vector. Where nucleic acids encoding both an elongase and acyltransferase enzymes are to be introduced, it may be preferable to incorporate both into a single vector. The vector may be, for example, a phage, plasmid or cosmid.

The vector will preferably also include one or more marker genes to enable the selection of plant cells which have been successfully transformed. Examples of suitable marker genes include antibiotic resistance genes such as those conferring resistance to kanamycin, G418 and hygromycin (*npt-II, hyg*-B); herbicide resistance genes such as those conferring resistance to phosphinothricin and sulfonamide based herbicides (*bar* and *sul* respectively; EP-A-242246 and EP-A-0369637) and screenable markers such as beta-glucoronidase (GB2197653), luciferase and green fluorescent protein.

The marker gene is preferably controlled by a second promoter which allows expression in cells other than the seed, thus allowing selection of cells or tissue containing the marker at any stage of development of the plant. Preferred second promoters are the promoter of nopaline synthase gene of *Agrobacterium* and the promoter derived from the gene which encodes the 35S subunit of cauliflower mosaic virus (CaMV) coat protein. However, any other suitable second promoter may be used.

In order to assist the process of deregulation, it is preferred that the selectable marker gene is not present in the plant which is to be grown commercially. Various techniques for marker elimination are available, including co-transformation followed by segregation and selection of segregants having the gene of interest but lacking the marker gene.

Methods for introducing nucleic acid into a plant are known to persons skilled in the art, and include techniques such as by the use of a disarmed Ti-plasmid vector carried by agrobacterium, for example as described in EP-A-0116718 and EP-A-0270822. Alternatively, the nucleic acid may be introduced by way of a particle gun, directly into the plant cells. This method is preferred for example where the plant is a monocot.

The oil may be extracted from the mature seeds of the plant by any suitable means which will be known to persons skilled in the art. For example, the methods described in Wilmer *et al* J Plant Physiol 147 486-492 (1996)) are available for laboratory scale extraction. For larger scale extraction, standard crushing techniques as practised by the industry may be used.

In a second aspect of the present invention, there is provided a plant capable of producing oil having an erucic acid content above 66%. Preferably, the plant is transgenic, and includes nucleic acid encoding an elongase enzyme capable of the production of very long chain fatty acids, including erucic acid, and nucleic acid encoding a 2-acyltransferase. Most preferably, the transgenes are in accordance with those defined in the first aspect of the invention.

The transgenic plants of the second aspect may be used in the production of tailored oils, which differ from native oils of the plant. In the present invention, the oil of the transgenic plant will differ from that of the native plant in terms of the erucic acid content of its oil. In particular, the erucic acid content of the oil will be higher than that of the native plant, and most preferably, it will be above 66%.

In a preferred embodiment of the invention, one or more of the plants' native oil biosynthesis enzymes may be rendered inoperative.

Any plant may be used in the present invention. Preferred plants are those whose seeds are used in the production of oil, for example *Brassica napus*, other *Brassica*, mustards, or other cruciferous plants, sunflower, soya or maize.

Also provided are plant parts, such as material required for propagating, in particular seeds.

A whole plant can be regenerated from a single transformed plant cell. Thus, in a

further aspect the present invention provides a transgenic plant cell including nucleic acid sequences encoding an elongase enzyme capable of the production of very long chain fatty acids, including erucic acid, and nucleic acid encoding a 2-acyltransferase. The regeneration can proceed by known methods (for *Brassica napus* see Moloney *et al* Plant Cell Reports 8 238-242, 1989).

In a third aspect of the invention, there is provided oil having an erucic acid content above 66%. Preferably, the oil is produced by a method of the first aspect. In a final aspect of the present invention, there is provided the use of such a transformed plant in the production of tailored oil.

Preferred features of the first aspect of the invention are for the other aspects *mutatis mutandis*.

The invention will now be described by way of non-limiting examples, with reference to the drawings in which:

Figure 1 shows the nucleic acid sequence encoding the *Brassica napus* enzyme FAE1-1.

Figure 2 shows the nucleic acid sequence encoding the *Brassica napus* enzyme FAE1-2.

Figure 3 shows the nucleic acid sequence encoding the 2-acyltransferase of *Limnanthes douglasii*.

Figure 4 shows the construction of plasmid pEW13.

Figure 5 shows the distribution of erucic acid levels, expressed as mol% of fatty acids, in seeds of LEAR transformants determined by gas chromatography (20:1 =

eicosenoic acid; 22:1 = erucic acid). Transgenic lines designated 333.n and 334.n were transformed with constructs containing the *fael-1* gene; lines designated 335.n and 336.n were transformed with constructs containing the *fael-2* gene . Data are derived from single samples of 20 seeds.

Figure 6 shows the distribution of erucic acid levels, expressed as mol% of fatty acids, in seeds of high erucic acid winter cultivar BGRV2 transformants determined by gas chromatography (20:1 = eicosenoic acid; 22:1 = erucic acid). Transgenic lines were transformed with either pEW13 (*fael-1* plus *lat2*) (Lines 337.n, 346.n, 349.n, 360.n, 398.n, 399.n) or pEW14 (*fael-2* plus *lat2*) (Lines 338.n, 347.n, 350.n, 361.n). Data are derived from single samples of 20 seeds.

Figure 7 shows the distribution of erucic acid content in selected progeny of primary transformed lines carrying the pEW14 transgenes.

a: suppressed BGRV2 line in T1; b: suppressed BGRV40 line in T1; c: very high-erucic acid BGRV2 in T1; d: very high-erucic acid BGRV40 in T1. The black bar in each panel represents the T1 parent seed, white bars represent the progeny seed.

EXAMPLES:

Example 1: Isolation of BnFAE1-1 and BnFAE1-2 and cloning into a binary vector.

Starting material

High erucic rape plants of a line BGRV2 were grown in the glasshouse to provide leaf material for DNA isolation. DNA was isolated essentially as described in Sambrook et al. (Molecular cloning: a laboratory manual, 2nd ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

PCR amplification of BnFAE1 sequences

Previously published *B.napus* fae1 sequences were analysed and two oligonucleotide primers were designed to amplify the *B.napus* fae1 genes from the high erucic acid winter line BGRV2:

BnFAE-F 5'CCTCATGACGTCCATTAACGTAAAGCTCC3' and

BnFAE-R 5'GTGAGCTCTTATTAGGACCGACCGTTTGGG3'.

PCR was performed using Tli Taq polymerase in the buffer supplied with1.5mM MgCl2, and an annealing temperature of 60°C. PCR products were cloned into pBluescriptII KS(+) (Stratagene) SmaI site and sequenced (MWG Biotech). (Figures 1 and 2)

The sequence of the BnFAE1-1 open reading frame shows higher homology to a published partial cDNA derived from *Brassica oleracea* than does BnFAE1-2. We therefore suggest that BnFAE1-1 represents the gene derived from the *B. oleracea* parent of *B. napus* (C-genome) and that BnFAE1-2 represents the *B. rapa* genome (A-genome).

Construction of plasmids pEW13 and pEW14

The two fae1 genes which had been amplified, cloned into pBluescript II, and sequenced were designated as pEW1 (fae1-1) and pEW3 (fae1-2). The fae1 genes were transferred as RcaI -EcoICRI fragments into pAR4 (Biogemma UK) NcoI-SmaI sites to place the genes between the Pnapin and CHS polyA sequences (pEW7 and 6). The two *fae1* expression cassettes were inserted adjacent to a similar Lat2 cassette, by digesting pEW7 and 6 with *SalI-SacI* and inserting the fragments into pT7Blue2 *SalI-SacI* sites (creating pEW9 and 8), before transfer of an *Eco*RI *Eco*ICRI fragment into the binary vector pSCVnos144 (Biogemma UK) *SalI-SmaI* sites resulting in pEW13 and14. The cloning strategy for the construction of pEW13 is shown in Figure 4.

Example 2: BnFAE1 genes encode elongase activities which specifically catalyse the formation erucic acid (22:1).

To demonstrate the product specificity of the two BnFAE1 genes, they were expressed in developing embryos of low erucic acid rape (LEAR).

Both plasmid pEW13 and plasmid pEW14 were transformed into Agrobacterium strain C58pMP90 and transformed into LEAR using agrobacterial transformation essentially as described in Moloney et al., (Plant Cell Reports 8: 238-242, 1989). Seed of the resulting transformed plants contain significant levels of erucic acid since both BnFAE1 genes complement mutations in the elongases of the LEAR line.

Oil extraction, separation, and analysis by GC was performed essentially as described in Wilmer et al (J. Plant Physiol. 147: 486-492, 1996). Oil was extracted from samples of 20 seed from each self-pollinated transformed plant. Seed were extacted with 1 ml chloroform, 2.8ml methanol/0.01 N hydrochloric acid (1:0.8, by vol.), containing 2 mg/ml triheptadecanoin as an internal standard. After shaking, 1 ml chloroform and 1 ml 0.01N hydrochloric acid were added and mixed. The phases were separated by

centrifugation in a bench-top centrifuge, and the aqueous phase washed with a further 1.5ml chloroform. The organic fraction contains most lipids.

Triglycerides were purified on 0.25 mm Silica gel 60 TLC plates, developed in hexane: diethyl ether: acetic acid (70:30:1 by vol.). After drying the lipids were visualised with iodine vapour and areas containing triglycerides were scraped off. TAG were transmethylated to produce fatty acid methyl esters (FAMEs) which could be analysed by gas chromatography.

FAME oil analysis was performed on mature seed of 11 plants transformed with constructs containing the *fael-1* gene and 25 plants transformed with constructs containing the *fael-2* gene demonstrating that the introduction of either *fael* gene led to a dramatic increase in the VLCFAs eicosenoic and erucic acid (C20:1 and C22:1) with a corresponding decrease in oleic acid (C18:1) (Figure 5). Both *fael* genes functioned well in the production of VLCFAs.

Example 3: BnFAE1 genes encode elongase activities which increase erucic acid levels in high-erucic acid oilseed rape.

A winter high-erucic acid line, BGRV2 was transformed with the two plasmids pEW13 and pEW14 previously tested in LEAR to examine whether the yield of erucic acid could be further enhanced in a cultivar which produced erucic acid.

Both plasmid pEW13 and plasmid pEW14 were transformed into Agrobacterium strain C58pMP90 and transformed into HEAR using agrobacterial transformation essentially as described in Moloney et al., (Plant Cell Reports 8: 238-242, 1989).

Oil analysis was performed on mature seed as described in Example 2.

Twenty seven plants were analysed (10 with *fae1-1* plus *lat2*, 17 with *fae1-2* plus *lat2*), 9 of which showed an increase in erucic acid of at least 4 mol%, the highest with 67.5 mol% fatty acids compared to the control of 54% as shown in Figure 6. Southern blot hybridisation of some of these plant lines showed that those with an increase in erucic acid had single or double insertions, whereas one of three lines which exhibited a considerable decrease (34%) contained 5 copies (data not shown), so that the phenotype observed could be due to co-suppression.

The data were applied to a t-test to determine whether there was a significant difference between the two populations of plants derived from *fae1-1* or *fae1-2* plus *lat2*. The probability value p = 0.00056917, demonstrated that there was a significant difference between the populations, and that *fae1-2* appeared to be more efficient.

The construct pEW14 *(fae1-2* plus *lat2)* was also used to transform a Canadian spring HEAR line BGRV40, with similar results. Three of the four transgenic plants produced, exhibited a considerable increase in erucic acid whereas one line was repressed.

In the lines derived from either of the HEAR varieties it is C22:1 which is specifically increased, not C20:1.

In all cases, it is the T1 generation that has been analysed, which is hemizygous. After self-fertilization and segregation analysis it will be possible to re-evaluate the erucic acid levels in the T2 generation, which may then exhibit further increases in C22:1 as the inserted T-DNA bearing the elongase and acyl-transferase cassette becomes homozygous.

Sn-2 analysis confirmed the incorporation of C22:1 at the Sn-2 position of the TAG by the 2-acyltransferase encoded by the introduced *lat2* gene.

Example 4: Very high erucic acid levels are stable in the next generation in oilseed rape.

Five lines of the winter high-erucic acid line, BGRV2, transformed with the plasmid pEW14, and three lines of the spring high-erucic acid line, BGRV40, transformed with the plasmid pEW14, chosen from the material described in Example 3 were grown in a following generation.

Four of the winter high erucic-acid lines and two of the spring high-erucic acid lines contained one or two copies of the transgene as determined by southern analysis, whereas one each of the winter and spring lines contained multiple copies of the transgenes.

Twenty plants per line were grown in the glasshouse and presence of the additional fae1-2 gene was confirmed using PCR with a primer sequence internal to the fae1-2 sequence and a primer inside the CHS terminator sequence. Mature seed was collected for analysis and oil analysis was performed on these seed as described in Example 2.

The two lines with multiple copies of the transgene that showed reduced levels of erucic acid in the T1 seed, both produced a wide range of erucic levels in the T2, as shown in figure 7.a and b. The lowest levels observed were below those observed in T1 seed, the highest levels well above the level of erucic acid observed in non-transgenic control oilseed rape of the appropriate line, and similar to the highest levels observed in the low-copy number lines, at least for the winter oilseed rape.

For the lines that showed increases in erucic acid in the T1 seed, this phenotype was maintained in the T2 seed, with the highest levels of erucic acid observed still around 67 to 68% (figure 7.c and d).

When the progeny of the lines with single or double copies of the transgene were tested by PCR for the presence of the transgene, segregation ratios of transgenic versus non-transgenic plants were observed that are consistent with a mendelian segregation of the transgene. Chi-square tests returned values showing non-significant deviation from an independent segregation model (data not shown).

CLAIMS

1. A method of producing, in a plant, oil having an erucic acid content above 66% erucic acid, the method comprising (i) expressing in the plant nucleic acid encoding an elongase and nucleic acid encoding an acyltransferase enzyme; and (ii) extracting oil from the plant.

2. A method according to claim 1, wherein the elongase is a fatty acid elongase enzyme from *Brassica napus,* or one substantially identical thereto.

3. A method according to claim 2, wherein the fatty acid elongase is encoded by the sequence of Figure 1 or 2, or a sequence substantially identical thereto.

4. A method according to any one of claims 1 to 3, wherein the acyltransferase is a 2-acyltransferase.

5. A method according to claim 4, wherein the 2-acyltransferase is from *L. douglassi.*

6. A method according to any one of claims 1 to 5, wherein the nucleic acid sequences are under the control of a regulatory sequence.

7. A method according to claim 6, wherein the regulatory sequence is an FAE promoter of *Brassica napus.*

8. A method according to any one of the previous claims wherein the nucleic acid sequences are provided in the form of a vector.

9. A method according to any one of the previous claims wherein the nucleic acid sequences are operably linked to one or more marker genes.

10. A method according to any one of the previous claims wherein the activity of the native elongase and/or acyltransferase of the plant is suppressed.

11. A method according to claim 10, wherein the activity is suppressed by antisense technology.

12. A transgenic plant capable of producing oil having an erucic acid content above 66%.

13. A transgenic plant according to claim 12 which expresses a 2-acyltransferase and an elongase enzyme which are not native to the plant.

14. A plant according to claim 12 or 13 which is *Brassica napus*.

15. A seed of a plant of any one of claims 12 to 14.

16. Oil produced by a method of any one of claims 1 to 11.

17. Oil extracted from a plant of any one of claims 12 to 14.

Figure 1

Sequence of the coding region of the Brassica napus FAE1-1 gene

```
1    ATGACGTCCA TTAACGTAAA GCTCCTTTAC CATTACGTCA TAACCAACCT
     TACTGCAGGT AATTGCATTT CGAGGAAATG GTAATGCAGT ATTGGTTGGA

51   TTTCAACCTT TGTTTCTTTC CATTAACGGC GATCGTCGCC GGAAAAGCCT
     AAAGTTGGAA ACAAAGAAAG GTAATTGCCG CTAGCAGCGG CCTTTTCGGA

101  ATCGGCTTAC CATAGACGAT CTTCACCACT TATACTATTC CTATCTCCAA
     TAGCCGAATG GTATCTGCTA GAAGTGGTGA ATATGATAAG GATAGAGGTT

151  CACAACCTCA TAACCATTGC TCCACTCTTT GCCTTCACCG TTTTCGGTTC
     GTGTTGGAGT ATTGGTAACG AGGTGAGAAA CGGAAGTGGC AAAAGCCAAG

201  GGTTCTCTAC ATCGCAACCC GGCCCAAACC GGTTTACCTC GTTGAGTACT
     CCAAGAGATG TAGCGTTGGG CCGGGTTTGG CCAAATGGAG CAACTCATGA

251  CATGCTACCT TCCACCAACG CATTGTAGAT CAAGTATCTC CAAGGTCATG
     GTACGATGGA AGGTGGTTGC GTAACATCTA GTTCATAGAG GTTCCAGTAC

301  GATATCTTTT ACCAAGTAAG AAAAGCTGAT CCTTCTCGGA ACGGCACGTG
     CTATAGAAAA TGGTTCATTC TTTTCGACTA GGAAGAGCCT TGCCGTGCAC

351  CGATGACTCG TCCTGGCTTG ACTTCTTGAG GAAGATTCAA GAACGTTCAG
     GCTACTGAGC AGGACCGAAC TGAAGAACTC CTTCTAAGTT CTTGCAAGTC

401  GTCTAGGCGA TGAAACCCAC GGGCCCGAGG GGCTGCTTCA GGTCCCTCCC
     CAGATCCGCT ACTTTGGGTG CCCGGGCTCC CCGACGAAGT CCAGGGAGGG

451  CGGAAGACTT TTGCGGCGGC GCGTGAAGAG ACGGAGCAAG TTATCATTGG
     GCCTTCTGAA AACGCCGCCG CGCACTTCTC TGCCTCGTTC AATAGTAACC

501  TGCGCTAGAA AATCTATTCA AGAACACCAA TGTTAACCCT AAAGATATAG
     ACGCGATCTT TTAGATAAGT TCTTGTGGTT ACAATTGGGA TTTCTATATC

551  GTATACTTGT GGTGAACTCA AGCATGTTTA ATCCAACTCC TTCGCTCTCC
     CATATGAACA CCACTTGAGT TCGTACAAAT TAGGTTGAGG AAGCGAGAGG

601  GCGATGGTCG TTAACACTTT CAAGCTCCGA AGCAACGTAA GAAGCTTTAA
     CGCTACCAGC AATTGTGAAA GTTCGAGGCT TCGTTGCATT CTTCGAAATT

651  CCTTGGTGGC ATGGGTTGTA GTGCCGGCGT TATAGCCATT GATCTAGCAA
     GGAACCACCG TACCCAACAT CACGGCCGCA ATATCGGTAA CTAGATCGTT

701  AGGACTTGTT GCATGTCCAT AAAAATACGT ATGCTCTTGT GGTGAGCACA
     TCCTGAACAA CGTACAGGTA TTTTTATGCA TACGAGAACA CCACTCGTGT

751  GAGAACATCA CTTATAACAT TTACGCTGGT GATAATAGGT CCATGATGGT
     CTCTTGTAGT GAATATTGTA AATGCGACCA CTATTATCCA GGTACTACCA

801  TTCAAATTGC TTGTTCCGTG TTGGTGGGGC CGCTATTTTG CTCTCCAACA
     AAGTTTAACG AACAAGGCAC AACCACCCCG GCGATAAAAC GAGAGGTTGT
```

```
      851  AGCCTAGAGA TCGTAGACGG TCCAAGTACG AGCTAGTTCA CACGGTTCGA
           TCGGATCTCT AGCATCTGCC AGGTTCATGC TCGATCAAGT GTGCCAAGCT

      901  ACGCATACCG GAGCTGACGA CAAGTCTTTT CGTTGCGTGC AACAAGGAGA
           TGCGTATGGC CTCGACTGCT GTTCAGAAAA GCAACGCACG TTGTTCCTCT

      951  CGATGAGAAC GGCAAAACCG GAGTGAGTTT GTCCAAGGAC ATAACCGATG
           GCTACTCTTG CCGTTTTGGC CTCACTCAAA CAGGTTCCTG TATTGGCTAC

     1001  TTGCTGGTCG AACGGTTAAG AAAAACATAG CAACGCTGGG TCCGTTGATT
           AACGACCAGC TTGCCAATTC TTTTTGTATC GTTGCGACCC AGGCAACTAA

     1051  CTTCCGTTAA GCGAGAAACT TCTTTTTTTC GTTACCTTCA TGGGCAAGAA
           GAAGGCAATT CGCTCTTTGA AGAAAAAAAG CAATGGAAGT ACCCGTTCTT

     1101  ACTTTTCAAA GACAAAATCA AACATTATTA CGTCCCGGAC TTCAAGCTTG
           TGAAAAGTTT CTGTTTTAGT TTGTAATAAT GCAGGGCCTG AAGTTCGAAC

     1151  CTATCGACCA TTTTTGTATA CATGCCGGAG GCAAAGCCGT GATTGATGTG
           GATAGCTGGT AAAAACATAT GTACGGCCTC CGTTTCGGCA CTAACTACAC

     1201  CTAGAGAAGA ACCTAGGCCT AGCACCGATC GATGTAGAGG CATCAAGATC
           GATCTCTTCT TGGATCCGGA TCGTGGCTAG CTACATCTCC GTAGTTCTAG

     1251  AACGTTACAT AGATTTGGAA ACACTTCATC TAGCTCAATA TGGTATGAGT
           TTGCAATGTA TCTAAACCTT TGTGAAGTAG ATCGAGTTAT ACCATACTCA

     1301  TGGCATACAT AGAAGCAAAA GGAAGGATGA AGAAAGGTAA TAAAGTTTGG
           ACCGTATGTA TCTTCGTTTT CCTTCCTACT TCTTTCCATT ATTTCAAACC

     1351  CAGATTGCTT TAGGGTCAGG CTTTAAGTGT AACAGTGCAG TTTGGGTGGC
           GTCTAACGAA ATCCCAGTCC GAAATTCACA TTGTCACGTC AAACCCACCG

     1401  TCTAAACAAT GTCAAAGCTT CAACAAATAG TCCTTGGGAA CACTGCATCG
           AGATTTGTTA CAGTTTCGAA GTTGTTTATC AGGAACCCTT GTGACGTAGC

     1451  ACAGATACCC GGTTAAAATT GATTCTGATT CAGGTAAGTC AGAGACTCGT
           TGTCTATGGG CCAATTTTAA CTAAGACTAA GTCCATTCAG TCTCTGAGCA

     1501  GTCCAAAACG GTCGGTCCTA A
           CAGGTTTTGC CAGCCAGGAT T
```

Figure 2

Sequence of the coding region of the Brassica napus FAE1-2 gene

```
   1  ATGACGTCCA TTAACGTAAA GCTCCTTTAC CATTACGTCA TAACCAACCT
      TACTGCAGGT AATTGCATTT CGAGGAAATG GTAATGCAGT ATTGGTTGGA

  51  TTTCAACCTT TGCTTCTTTC CGTTAACGGC GATCGTCGCC GGAAAAGCCT
      AAAGTTGGAA ACGAAGAAAG GCAATTGCCG CTAGCAGCGG CCTTTTCGGA

 101  ATCGGCTTAC CATAGACGAT CTTCACCACT TATACTATTC CTATCTCCAA
      TAGCCGAATG GTATCTGCTA GAAGTGGTGA ATATGATAAG GATAGAGGTT

 151  CACAACCTCA TAACCATCGC TCCACTCTTT GCCTTCACCG TTTTCGGTTC
      GTGTTGGAGT ATTGGTAGCG AGGTGAGAAA CGGAAGTGGC AAAAGCCAAG

 201  GGTTCTCTAC ATCGCAACCC GGCCCAAACC GGTTTACCTC GTTGAGTACT
      CCAAGAGATG TAGCGTTGGG CCGGGTTTGG CCAAATGGAG CAACTCATGA

 251  CATGCTACCT TCCACCAACG CATTGTAGAT CAAGTATCTC CAAGGTCATG
      GTACGATGGA AGGTGGTTGC GTAACATCTA GTTCATAGAG GTTCCAGTAC

 301  GATATCTTTT ATCAAGTAAG AAAAGCTGAT CCTTCTCGGA ACGGCACGTG
      CTATAGAAAA TAGTTCATTC TTTTCGACTA GGAAGAGCCT TGCCGTGCAC

 351  CGATGACTCG TCGTGGCTTG ACTTCTTGAG GAAGATTCAA GAACGTTCAG
      GCTACTGAGC AGCACCGAAC TGAAGAACTC CTTCTAAGTT CTTGCAAGTC

 401  GTCTAGGCGA TGAAACTCAC GGGCCCGAGG GGCTGCTTCA GGTCCCTCCC
      CAGATCCGCT ACTTTGAGTG CCCGGGCTCC CCGACGAAGT CCAGGGAGGG

 451  CGGAAGACTT TTGCGGCGGC GCGTGAAGAG ACGGAGCAAG TTATCATTGG
      GCCTTCTGAA AACGCCGCCG CGCACTTCTC TGCCTCGTTC AATAGTAACC

 501  TGCACTAGAA AATCTATTCA AGAACACCAA CGTTAACCCT AAAGATATAG
      ACGTGATCTT TTAGATAAGT TCTTGTGGTT GCAATTGGGA TTTCTATATC

 551  GTATACTTGT GGTGAACTCA AGCATGTTTA ATCCAACTCC ATCGCTCTCC
      CATATGAACA CCACTTGAGT TCGTACAAAT TAGGTTGAGG TAGCGAGAGG

 601  GCGATGGTCG TTAACACTTT CAAGCTCCGA AGCAACGTAA GAAGCTTTAA
      CGCTACCAGC AATTGTGAAA GTTCGAGGCT TCGTTGCATT CTTCGAAATT

 651  CCTTGGTGGC ATGGGTTGTA GTGCCGGCGT TATAGCCATT GATCTAGCAA
      GGAACCACCG TACCCAACAT CACGGCCGCA ATATCGGTAA CTAGATCGTT

 701  AGGACTTGTT GCATGTCCAT AAAAATACGT ATGCTCTTGT GGTGAGCACA
      TCCTGAACAA CGTACAGGTA TTTTTATGCA TACGAGAACA CCACTCGTGT

 751  GAGAACATCA CTTATAACAT TTACGCTGGT GATAATAGGT CCATGATGGT
      CTCTTGTAGT GAATATTGTA AATGCGACCA CTATTATCCA GGTACTACCA

 801  TTCAAATTGC TTGTTCCGTG TTGGTGGGGC CGCTATTTTG CTCTCCAACA
      AAGTTTAACG AACAAGGCAC AACCACCCCG GCGATAAAAC GAGAGGTTGT
```

```
851   AGCCTGGAGA TCGTAGACGG TCCAAGTACG AGCTAGTTCA CACGGTTCGA
      TCGGACCTCT AGCATCTGCC AGGTTCATGC TCGATCAAGT GTGCCAAGCT

901   ACGCATACCG GAGCTGACGA CAAGTCTTTT CGTTGCGTGC AACAAGGAGA
      TGCGTATGGC CTCGACTGCT GTTCAGAAAA GCAACGCACG TTGTTCCTCT

951   CGATGAGAAC GGCAAAATCG GAGTGAGTTT GTCCAAGGAC ATAACCGATG
      GCTACTCTTG CCGTTTTAGC CTCACTCAAA CAGGTTCCTG TATTGGCTAC

1001  TTGCTGGTCG AACGGTTAAG AAAAACATAG CAACGTTGGG TCCGTTGATT
      AACGACCAGC TTGCCAATTC TTTTTGTATC GTTGCAACCC AGGCAACTAA

1051  CTTCCGTTAA GCGAGAAACT TCTTTTTTTC GTTACCTTCA TGGGCAAGAA
      GAAGGCAATT CGCTCTTTGA AGAAAAAAAG CAATGGAAGT ACCCGTTCTT

1101  ACTTTTCAAA GATAAAATCA AACATTACTA CGTCCCGGAT TTCAAACTTG
      TGAAAAGTTT CTATTTTAGT TTGTAATGAT GCAGGGCCTA AAGTTTGAAC

1151  CTATTGACCA TTTTTGTATA CATGCCGGAG GCAGAGCCGT GATTGATGTG
      GATAACTGGT AAAAACATAT GTACGGCCTC CGTCTCGGCA CTAACTACAC

1201  CTAGAGAAGA ACCTAGCCCT AGCACCGATC GATGTAGAGG CATCAAGATC
      GATCTCTTCT TGGATCGGGA TCGTGGCTAG CTACATCTCC GTAGTTCTAG

1251  AACGTTACAT AGATTTGGAA ACACTTCATC TAGCTCAATA TGGTATGAGT
      TTGCAATGTA TCTAAACCTT TGTGAAGTAG ATCGAGTTAT ACCATACTCA

1301  TGGCATACAT AGAAGCAAAA GGAAGGATGA AGAAAGGTAA TAAAGTTTGG
      ACCGTATGTA TCTTCGTTTT CCTTCCTACT TCTTTCCATT ATTTCAAACC

1351  CAGATTGCTT TAGGGTCAGG CTTTAAGTGT AACAGTGCAG TTTGGGTGGC
      GTCTAACGAA ATCCCAGTCC GAAATTCACA TTGTCACGTC AAACCCACCG

1401  TCTAAACAAT GTCAAAGCTT CGACAAATAG TCCTTGGGAA CACTGCATCG
      AGATTTGTTA CAGTTTCGAA GCTGTTTATC AGGAACCCTT GTGACGTAGC

1451  ACAGATACCC GGTCAAAATT GATTCTGATT CAGGTAAGTC AGAGACTCGT
      TGTCTATGGG CCAGTTTTAA CTAAGACTAA GTCCATTCAG TCTCTGAGCA

1501  GTCCAAAACG GTCGGTCCTA A
      CAGGTTTTGC CAGCCAGGAT T
```

Figure 3

Sequence of the coding region of the LPAAT gene LAT2 from Limnanthes douglasii.

```
   1  ATGGCCAAAA CTAGAACTAG CTCTCTCCGC AACAGGAGAC AACTAAAGCC
      TACCGGTTTT GATCTTGATC GAGAGAGGCG TTGTCCTCTG TTGATTTCGG

  51  GGCTGTAGCT GCTACTGCTG ATGATGATAA AGATGGGGTT TTTATGGTAT
      CCGACATCGA CGATGACGAC TACTACTATT TCTACCCCAA AAATACCATA

 101  TGCTATCGTG TTTTAAAATT TTTGTTTGCT TTGCCATAGT GTTGATCACC
      ACGATAGCAC AAAATTTTAA AAACAAACGA AACGGTATCA CAACTAGTGG

 151  GCGGTGGCAT GGGGACTAAT CATGGTCTTG CTCTTACCTT GGCCTTATAT
      CGCCACCGTA CCCCTGATTA GTACCAGAAC GAGAATGGAA CCGGAATATA

 201  GAGGATTCGA CTAGGAAATC TATACGGCCA TATCATTGGT GGATTAGTGA
      CTCCTAAGCT GATCCTTTAG ATATGCCGGT ATAGTAACCA CCTAATCACT

 251  TATGGATTTA CGGAATACCA ATAAAGATCC AAGGATCCGA GCATACAAAG
      ATACCTAAAT GCCTTATGGT TATTTCTAGG TTCCTAGGCT CGTATGTTTC

 301  AAGAGGGCCA TTTATATAAG CAATCATGCA TCTCCTATCG ATGCTTTCTT
      TTCTCCCGGT AAATATATTC GTTAGTACGT AGAGGATAGC TACGAAAGAA

 351  TGTTATGTGG TTGGCTCCCA TAGGCACAGT TGGTGTTGCA AAGAAAGAGG
      ACAATACACC AACCGAGGGT ATCCGTGTCA ACCACAACGT TTCTTTCTCC

 401  TTATATGGTA TCCGCTGCTT GGACAACTAT ATACATTAGC CCATCATATT
      AATATACCAT AGGCGACGAA CCTGTTGATA TATGTAATCG GGTAGTATAA

 451  CGCATAGATC GGTCAAACCC GGCTGCGGCT ATTCAGTCTA TGAAAGAGGC
      GCGTATCTAG CCAGTTTGGG CCGACGCCGA TAAGTCAGAT ACTTTCTCCG

 501  AGTTCGTGTA ATAACCGAAA AGAATCTCTC TCTGATTATG TTTCCAGAGG
      TCAAGCACAT TATTGGCTTT TCTTAGAGAG AGACTAATAC AAAGGTCTCC

 551  GAACCAGGTC GAGAGATGGG CGTTTACTTC CTTTCAAGAA GGGTTTTGTT
      CTTGGTCCAG CTCTCTACCC GCAAATGAAG GAAAGTTCTT CCCAAAACAA

 601  CATCTAGCAC TTCAGTCACA TCTCCCAATA GTTCCGATGA TCCTTACAGG
      GTAGATCGTG AAGTCAGTGT AGAGGGTTAT CAAGGCTACT AGGAATGTCC

 651  TACACATTTA GCATGGAGGA AAGGTACCTT CCGTGTCCGG CCAGTACCCA
      ATGTGTAAAT CGTACCTCCT TTCCATGGAA GGCACAGGCC GGTCATGGGT

 701  TCACTGTCAA GTACCTTCCT CCTATAAACA CTGATGATTG GACTGTTGAC
      AGTGACAGTT CATGGAAGGA GGATATTTGT GACTACTAAC CTGACAACTG

 751  AAAATCGACG ATTACGTCAA AATGATACAC GACGTCTATG TCCGCAACCT
      TTTTAGCTGC TAATGCAGTT TTACTATGTG CTGCAGATAC AGGCGTTGGA

 801  ACCTGCGTCT CAAAAACCAC TTGGTAGCAC AAATCGCTCA AAT
      TGGACGCAGA GTTTTTGGTG AACCATCGTG TTTAGCGAGT TTA
```

Figure 4
Construction of pEW13

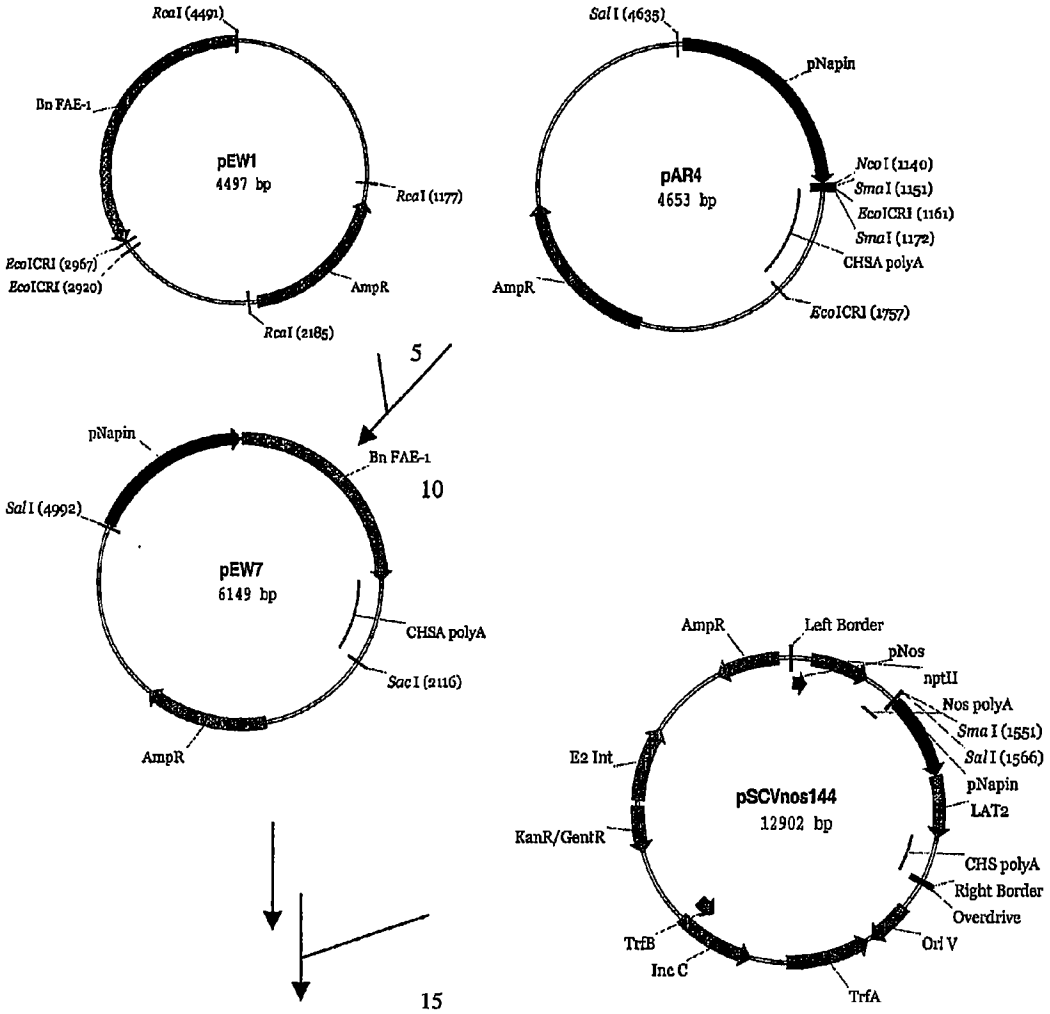
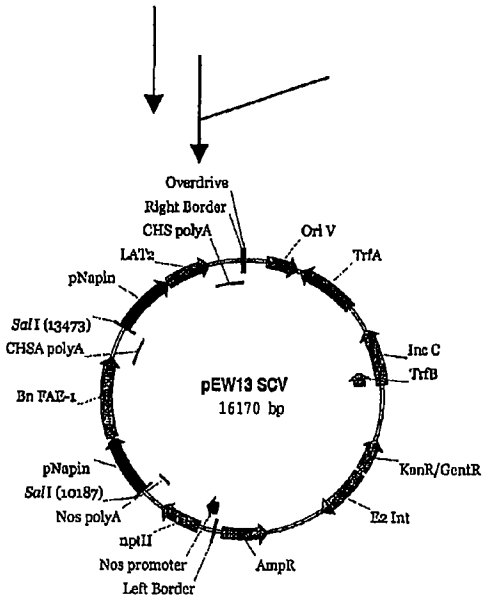



Figure 4 (continued)